
 W P S E R E (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_PP protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 08:46:40 2000; MasPar time 4.48 Seconds
 Tabular output not generated. 317.408 Million cell updates/sec

Title: >US-09-331-631-22
 Description: (23-84) from US09331631.pep
 Perfect Score: 459
 Sequence: 1 EDDNHHHGHGKSGQCVRR.....EKROERSRHEADRSGEGSS 60

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs, 23666106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:geneseqp

Statistics: Mean 24.771; Variance 107.255; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	459	100.0	593	1	Zea mays antimicrobial	8.76e-34
2	180	39.2	637	1	Hordeum vulgare antim	1.10e-07
3	121	26.4	625	1	Macadamia integrifolia	1.48e-02
4	115	25.1	666	1	Macadamia integrifolia	4.69e-02
5	109	23.7	666	1	Macadamia integrifolia	1.47e-01
6	102	22.2	432	1	Human regulatory molec	5.49e-01
7	96	20.9	590	1	Gossypium hirsutum ant	1.67e-00
8	91	19.8	33	1	Zea mays antimicrobial	4.18e-00
9	91	19.8	450	1	Tyrosine kinase associ	4.18e-00
10	89	19.4	186	1	Trypanosoma cruzi anti	6.00e-00
11	88	19.2	388	1	Mycobacterium spec	7.19e-00
12	87	19.0	355	1	Human clone 56 protein	8.61e-00
13	87	19.0	524	1	Miniature swine retrov	1.23e-01
14	86	18.7	434	1	Peptide fragment of N-	1.03e-01
15	86	18.7	567	1	Polypeptide fragment e	1.03e-01
16	86	18.7	797	1	Murine PgC-1 protein.	1.03e-01
17	85	18.5	314	1	Human stomach carcino	1.23e-01
18	85	18.5	525	1	Theobroma cacao anti	1.23e-01
19	85	18.5	566	1	Sequence encoded by 67	1.23e-01
20	82	17.9	660	1	B subunit of human fac	2.10e-01
21	82	17.9	898	1	Mycobacterium tubercul	2.10e-01
22	81	17.6	444	1	G. max truncated SBP1	2.51e-01
23	81	17.6	524	1	G. max SBP1 protein.	2.51e-01

ID	Score	Query Match	Length	ID	Description	Pred. No.
24	81	17.6	1148	1	Renal cancer associate	2.51e-01
25	81	17.6	2337	1	Human calcium channel	2.51e-01
26	80	17.4	147	1	Mycobacterium species	2.99e-01
27	80	17.4	233	1	Mycobacterium species	2.99e-01
28	80	17.4	288	1	Human cardiac tropoin	2.99e-01
29	80	17.4	288	1	Recombinant human card	2.99e-01
30	80	17.4	409	1	G. max truncated SBP2	2.99e-01
31	80	17.4	436	1	Human 70K UI snRNP pro	2.99e-01
32	80	17.4	489	1	G. max SBP2 protein.	2.99e-01
33	80	17.4	614	1	70K autoantigen, part	2.99e-01
34	80	17.4	1311	1	Precis coenia patched	2.99e-01
35	80	17.4	1311	1	Precis coenia (butterf	2.99e-01
36	80	17.4	2237	1	Sequence of the alpha	2.99e-01
37	79	17.2	303	1	Dirofilaria immitis pa	3.56e-01
38	79	17.2	467	1	Rat Glucocorticoid ind	3.56e-01
39	79	17.2	605	1	Human neuronal calcium	3.56e-01
40	79	17.2	612	1	Human neuronal calcium	3.56e-01
41	79	17.2	660	1	Human neuronal calcium	3.56e-01
42	79	17.2	1382	1	Human metastasis-assoc	3.56e-01
43	79	17.2	1829	1	Actin-filament binding	3.56e-01
44	78	17.0	162	1	Sequence of a new cyto	4.24e-01
45	78	17.0	2339	1	Human calcium channel	4.24e-01

ALIGNMENTS

ID	Score	Query Match	Length	ID	Description	Pred. No.
1	459	100.0%	593	1	Zea mays antimicrobial	8.76e-34
2	180	39.2%	637	1	Hordeum vulgare antim	1.10e-07
3	121	26.4%	625	1	Macadamia integrifolia	1.48e-02
4	115	25.1%	666	1	Macadamia integrifolia	4.69e-02
5	109	23.7%	666	1	Macadamia integrifolia	1.47e-01
6	102	22.2%	432	1	Human regulatory molec	5.49e-01
7	96	20.9%	590	1	Gossypium hirsutum ant	1.67e-00
8	91	19.8%	33	1	Zea mays antimicrobial	4.18e-00
9	91	19.8%	450	1	Tyrosine kinase associ	4.18e-00
10	89	19.4%	186	1	Trypanosoma cruzi anti	6.00e-00
11	88	19.2%	388	1	Mycobacterium spec	7.19e-00
12	87	19.0%	355	1	Human clone 56 protein	8.61e-00
13	87	19.0%	524	1	Miniature swine retrov	1.23e-01
14	86	18.7%	434	1	Peptide fragment of N-	1.03e-01
15	86	18.7%	567	1	Polypeptide fragment e	1.03e-01
16	86	18.7%	797	1	Murine PgC-1 protein.	1.03e-01
17	85	18.5%	314	1	Human stomach carcino	1.23e-01
18	85	18.5%	525	1	Theobroma cacao anti	1.23e-01
19	85	18.5%	566	1	Sequence encoded by 67	1.23e-01
20	82	17.9%	660	1	B subunit of human fac	2.10e-01
21	82	17.9%	898	1	Mycobacterium tubercul	2.10e-01
22	81	17.6%	444	1	G. max truncated SBP1	2.51e-01
23	81	17.6%	524	1	G. max SBP1 protein.	2.51e-01

PS Claim 1: Page 60-62; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 637 AA;

Query Match 39.2%; Score 180; DB 1; Length 637;
 Best Local Similarity 33.9%; Pred. No. 1.10e-07;
 Matches 19; Conservative 22; Mismatches 13; Indels 2; Gaps 2;
 Db 30 DDEDDRGRGHSLQOCVORCRORPRYSARCAVOCRDQOQHGHEDEEGRG 85
 Y 26 DDNNHHHGHGKSGQCVARC-EDRPWHRPRLCEGCREREK-RQESRHEADRS 79

RESULT 3
 ID W62830 standard; Protein; 625 AA.
 AC W62830;
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT /note="signal peptide"
 FT Protein 29..666
 FT /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB; V42316.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling macrobrial infestations of plants or mammals
 PS Claim 1: Page 43-45; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 625 AA;

Query Match 26.4%; Score 121; DB 1; Length 625;
 Best Local Similarity 31.7%; Pred. No. 1.48e-02;
 Matches 19; Conservative 15; Mismatches 23; Indels 3; Gaps 3;
 Db 73 EFNRRDPOQOYECCRCORRETEPRHMQICOCRCRYEKRKQKRYEEOHREDE 132
 Y 25 EDDNNHHHGHGKSGQCVARCEDRPWHRP-R-CLEOC-REEREKROERSHHEADRSGE 81

RESULT 4
 ID W62828 standard; Protein; 666 AA.
 AC W62828;
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT /note="signal peptide"
 FT Protein 29..666
 FT /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB; V42310.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -

PT useful for controlling macrobrial infestations of plants or mammals
 PS Claim 1: Page 34-36; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 666 AA;

Query Match 25.1%; Score 115; DB 1; Length 666;
 Best Local Similarity 35.7%; Pred. No. 4.69e-02;
 Matches 15; Conservative 14; Mismatches 11; Indels 2; Gaps 2;
 Db 182 EEDNRDPOQREYEDRCRCRCEOEPRQOH-OCOLRCREOQRO 222
 Y 25 EDDNNHHHGHGKSGQCVARCEDR-PWHRPRLCEGCRERE 65

RESULT 5
 ID W62829 standard; Protein; 666 AA.
 AC W62829;
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT /note="signal peptide"
 FT Protein 29..666
 FT /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997; AU0874.
 PR 20-DEC-1996; AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NT, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-377279/32.
 DR N-PSDB; V42311.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling macrobrial infestations of plants or mammals
 PS Claim 1: Page 39-41; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 666 AA;

Query Match 23.7%; Score 109; DB 1; Length 666;
 Best Local Similarity 31.1%; Pred. No. 1.47e-01;
 Matches 19; Conservative 16; Mismatches 21; Indels 5; Gaps 4;
 Db 114 EFNRRDPOQOYECCRCORRETEPRHMQICOCRCRYEKRKQKRYEEOHREDE 172
 Y 25 EDDNNHHHGHGKSGQCVARCE--DR-PWHRPRLCEOC-REEREKROERSHHEADRS 80

RESULT 6
 ID W93954 standard; Protein; 432 AA.
 AC W93954;
 DT 30-JUN-1999 (first entry)
 DE Human regulatory molecule,HRM-10 protein.
 KW Human regulatory molecule; HRM-10; cytoskeletal activity; immune modulator;
 KW transcription factor; enhancer; cell proliferation stimulation; cancer;
 KW treatment; microarray; detection; diagnosis; cell proliferation disease;
 KW leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone;
 KW brain; lung; liver; ovary; skin; teratocarcinoma; immune response;
 KW allergy; asthma; diabetes; multiple sclerosis; Grave's disease;
 OS Homo sapiens.
 PN W09915658-A2.
 PD 01-APR-1999.
 PF 22-SEP-1998; U19839.


```

RESULT 10
ID W26536 standard; Protein: 186 AA.
AC W26536.
DE 12-JAN-1998 (first entry)
DE Trypanosoma cruzi antigen.
KW Antigen; epitope; vaccine; protective immunity; Chagas disease;
KW diagnosis; therapy; immunosassay.
OS Trypanosoma cruzi Tulahen strain C2.
PN W09718475-A1.
PD 22-MAY-1997.
PF 14-NOV-1996; U18624.
PR 14-NOV-1995; US-557309.
PA (CORI-) CORIXA CORP.
PI Houghton RL, Lodes MJ, Reed SG, Skeiky YAM;
PI WPI: 97-289413/26.
DE Diagnosing Trypanosoma cruzi infection by detecting antibodies to
PT novel antigens - which are useful in vaccines to provide protective
PS immunity against Chagas' disease
PS Disclosure: Page 70-71; 110pp; English.
CC This polypeptide sequence comprises a Trypanosoma cruzi full-length
CC antigen. T. cruzi antigens (see W26530-41), and epitope-containing
CC repeat sequences (see W19094-102, W19079-86 and W26542-44) of
CC native T. cruzi antigens, have been identified by sequencing
CC genomic DNA and cDNA clones (see T69151-72) isolated from T. cruzi
CC expression libraries by screening with sera from infected patients.
CC The antigens can be expressed in transformed or transfected host
CC cells. They can be used in a variety of immunassays for detecting
CC T. cruzi infection in a blood, serum, plasma, saliva, cerebrospinal
CC fluid or urine sample. They are also useful in vaccines and
CC pharmaceutical compositions for inducing protective immunity against
CC Chagas disease.
CC Sequence 186 AA;

Query Match 19.4%; Score 89; DB 1; Length 186;
Best Local Similarity 32.5%; Pred. No. 6.00e+00;
Matches 13; Conservative 14; Mismatches 11; Indels 2; Gaps 2;

Db 128 RREAFERAREVERAROEAEELAROSEERARORAEERA 167
OY 42 RR-CEDRPMHQ-RPRCLDCEEREERERKERSRHEADRS 79

RESULT 11
ID Y04998 standard; Protein: 388 AA.
AC Y04998.
DE 06-JUL-1999 (first entry)
DE Mycobacterium species protein sequence 50B.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection.
OS Mycobacterium sp.
PN W09909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pellicic V, Portnoi D, Gognet de la Salmoniere Y,
PI Guigueno A;
PI WPI: 99-181045/15.
DE Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 32; Fig 50B; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
CC Sequence 388 AA;

```

```

Query Match 19.2%; Score 88; DB 1; Length 388;
Best Local Similarity 35.0%; Pred. No. 7.19e+00;
Matches 14; Conservative 10; Mismatches 14; Indels 2; Gaps 2;

Db 37 RRCQROPRWRMPRCRWPMALRRGRTRTRRNP 76
OY 42 RCEDRP-WHORPR-CLEOCREEREERKERSRHEADRS 79

RESULT 12
ID W05398 standard; Protein: 355 AA.
AC W05398.
DE 19-FEB-1998 (first entry)
DE Human clone 56 protein.
KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW cellular signalling element; cellular structural element; malignancy;
KW protein identification; functional domain; protein screening;
OS Homo sapiens.
PN W09631625-A1.
PD 10-OCT-1996.
PF 04-APR-1996; U04454.
PR 03-APR-1996; US-630915.
PR 07-APR-1995; US-417872.
PA (CYTO-) CYTOGEN CORP.
PA (UYN-) UNIV NORTH CAROLINA.
PI Foxlkes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;
PI WPI: 96-465045/46.
DE Identifying polypeptide(s) having specific functional domain (esp.
PT SH3 domain) - comprises detecting selective binding to recognition
PT unit, regardless of sequence homology
PS Claim 54; Fig 57; 174pp; English.
CC W05386-W05403 represent novel human and mouse Src-homology region 3 (SH3)
CC domain containing proteins that can be used in the method of the
CC invention. SH3 domain containing proteins play a role in signalling and
CC structural elements of cells. The method of the invention is for
CC identifying polypeptides containing functional domains of interest
CC (especially SH3 domains). The method comprises contacting a multivalent
CC recognition unit (RU) complex with a number of peptides and identifying
CC polypeptides having a selective binding affinity for the RU complex. The
CC method is based on functional similarities and does not rely on sequence
CC similarities. Prior methods only gave limited success for identifying
CC proteins which contain an SH3 domain due to the minimal sequence homology
CC among known SH3 proteins. It has been found that small peptide RUS in
CC multivalent form have reduced specificity for a given functional domain
CC compared to monomer RUS. Multivalent RU complexes are particularly suited
CC to screening for polypeptides containing functional domains that are
CC similar to, but not identical in sequence to, the original target
CC functional domain. The new method enables proteins having a common
CC function to be identified. Identification of novel SH3 proteins will be
CC useful for a better understanding of cell growth, malignancy, signal
CC transduction processes, etc. New candidate drugs can be identified, and
CC their specificities (e.g. pharmacological activities) can be assessed
CC using the method of the invention.
CC Sequence 355 AA;

Query Match 19.0%; Score 87; DB 1; Length 355;
Best Local Similarity 37.8%; Pred. No. 8.61e+00;
Matches 17; Conservative 11; Mismatches 14; Indels 3; Gaps 3;

Db 53 RLUEKRAEBAQRLQLEERRE-REUEARARORGYOEGGEASP 96
OY 42 RCEDRP-WHORPR-CLEOCREEREERKERSRHEADRS 84

RESULT 13
ID W32096 standard; Protein: 524 AA.
AC W32096.
DE 09-FEB-1998 (first entry)
DE Miniature swine retrovirus GAG protein.
KW Retrovirus; porcine; GAG protein; xenotransplantation; infectious;

```


PR 06-JUN-1997: US-048970.
PR 06-JUN-1997: US-048974.
PR 06-JUN-1997: US-049373.
PR 05-SEP-1997: US-057384.
PR 05-SEP-1997: US-057629.
PR 05-SEP-1997: US-057642.
PR 05-SEP-1997: US-057645.
PR 05-SEP-1997: US-057648.
PR 05-SEP-1997: US-057651.
PR 05-SEP-1997: US-057662.
PR 05-SEP-1997: US-057668.
PR 05-SEP-1997: US-057762.
PR 05-SEP-1997: US-057765.
PR 05-SEP-1997: US-057771.
PR 05-SEP-1997: US-057776.
PR 06-JUN-1997: US-048876.
PR 06-JUN-1997: US-048880.
PR 06-JUN-1997: US-048883.
PR 06-JUN-1997: US-048892.
PR 06-JUN-1997: US-048895.
PR 06-JUN-1997: US-048898.
PR 06-JUN-1997: US-048901.
PR 06-JUN-1997: US-048917.
PR 06-JUN-1997: US-048963.
PR 06-JUN-1997: US-048971.
PR 06-JUN-1997: US-049019.
PR 06-JUN-1997: US-049374.
PR 05-SEP-1997: US-057627.
PR 05-SEP-1997: US-057634.
PR 05-SEP-1997: US-057643.
PR 05-SEP-1997: US-057646.
PR 05-SEP-1997: US-057649.
PR 05-SEP-1997: US-057654.
PR 05-SEP-1997: US-057656.
PR 05-SEP-1997: US-057660.
PR 05-SEP-1997: US-057763.
PR 05-SEP-1997: US-057769.
PR 05-SEP-1997: US-057774.
PR 05-SEP-1997: US-057777.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,
PI Florence K, Greene JW, Hu J, Kraw H, Lafleur DW,
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z:
DR WPI: 99-059865/05.
DR N-PsDB: V84446.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Disclosure: Page 30: 772PP: English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate cancers, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The